# Quasi-Concavity for Gaussian Multicast Relay Channels

Mohit Thakur and Gerhard Kramer



***Abstract*—Upper and lower bounds on the capacity of Gaussian multicast relay channels are shown to be quasi-concave in the receiver signal-to-noise ratios and the transmit correlation coefficient. The bounds considered are the cut-set bound, decode-forward (DF) rates, and quantize-forward rates. The DF rates are shown to be quasi-concave in the relay position and this property is used to optimize the relay position for several networks.**

***Index Terms*—capacity, decode-forward, multicast, relaying**

## I. Introduction

A *multicast relay channel* (MRC) is an information network with a source node, a relay node, and two or more destination nodes, and where one message originating at the source should be received reliably at the destinations. We consider additive white Gaussian noise (AWGN) MRCs and show that certain information rate expressions are quasi-concave in the receiver signal-to-noise ratios (SNRs), the transmit correlation coefficient, and the relay position. Quasi-concavity means that efficient algorithms can optimize signaling and the relay position. For example, suppose one wishes to place a relay to maximize the multicast rate [1], [2]. We studied this problem for a decode-forward (DF) strategy in the low-SNR regime in [3]–[5]. This paper extends the results to other strategies such as quantize-forward (QF) and to general SNR regimes.

The paper is organized as follows. Section II presents the MRC model and bounds on the MRC capacity. Sec. III shows that these bounds are quasi-concave in a correlation coefficient and the SNRs. Sec. IV shows that the best DF rate is quasi-concave in the relay position. Sec. V compares the performance of different DF strategies, and computes optimal relay positions. Sec. VI discusses complex channel models and Sec. VII concludes the paper. The Appendices provide useful results on concavity and quasi-concavity.

## II. MRC Model and Information Rates

### A. Model

An MRC has three types of nodes:

- a source node $s$ that generates a message $W$ and transmits the symbols $X_s^n = X_{s,1}, X_{s,2}, \ldots, X_{s,n}$;

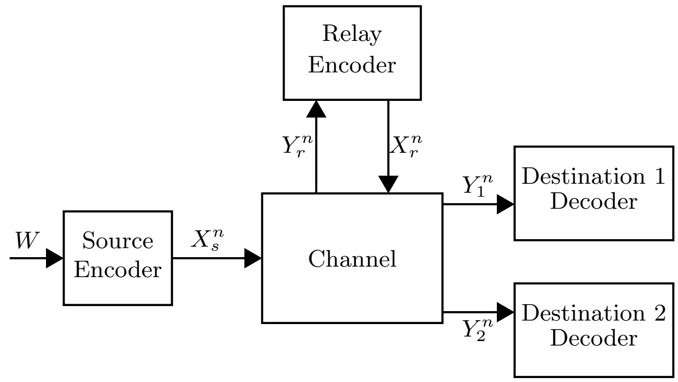


Fig. 1: MRC with two destinations.

- a relay node $r$ that receives and forwards symbols $Y_{r,k}$ and $X_{r,k}$, respectively, for $k = 1, 2, \cdots, n$;
- destination nodes $j = 1, 2, \ldots, N$ where node $j$ receives $Y_j^n = Y_{j,1}, Y_{j,2}, \ldots, Y_{j,n}$ and estimates $W$ as $\hat{W}_j$.

We denote the destination node set as $\mathcal{T} = \{1, 2, \ldots, N\}$. The classic relay channel has $N = 1$ and Fig. 1 shows an MRC with $N = 2$. A *memoryless* MRC has a function $h(\cdot)$ and a noise random variable $\mathbf{Z}$ so that for every time instant the $N+1$ channel outputs $\mathbf{Y} = (Y_r\ Y_1\ \ldots\ Y_N)$ are given by

$$\mathbf{Y} = h(X_s, X_r, \mathbf{Z}).$$

The noise $\mathbf{Z}$ is statistically independent of $X_s$ and $X_r$, and the noise variables at different times are statistically independent.

An encoding strategy has

- a message $W$ uniformly distributed over $\{1, 2, \ldots, M\}$;
- an encoding function $e_s$ such that $X_s^n = e_s(W)$;
- relay functions $e_{r,k}$ with $X_{r,k} = e_{r,k}(Y_{r,1}, .., Y_{r,k-1})$, where $k = 1, 2, \ldots, n$;
- decoding functions $d_j$ such that $d_j(Y_j^n) = \hat{W}_j$, $j \in \mathcal{T}$.

The error probability at destination $j$ is $P_{e,j} = \Pr\left[\hat{W}_j \neq W\right]$. The multicast rate is $R = (\log_2 M)/n$ bits/use. The rate $R$ is *achievable* if, for any $\epsilon > 0$ and sufficiently large $n$, there is an encoding strategy with $P_{e,j} \leq \epsilon$ for all $j \in \mathcal{T}$. The *capacity* $C$ is the supremum of the achievable rates.

### B. Information Rates

The following bounds were given in [6] for the relay channel ($N = 1$). Their extensions to MRCs are straightforward.

- Cut-Set Bound: $C \leq R_{CS}$ where

$$R_{CS} = \max\left\{\min_{1\leq j\leq N} \min(I(X_sX_r;Y_j), I(X_s;Y_rY_j|X_r))\right\} \quad (1)$$

Date of current version June 10, 2015. Submitted for publication in the IEEE Transactions on Wireless Communications, June 2015.

M. Thakur and G. Kramer were supported by an Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research. The paper was presented in part at the 2015 IEEE International Symposium on Information Theory, Hong Kong.

M. Thakur was with the Institute for Communications Engineering, Technische Universität München, 80333 Munich, Germany. He is now with the Institute of Communications and Navigation at the German Aerospace Center (DLR), 82234 Weßling, Germany (e-mail: mohit.thakur@dlr.de). G. Kramer is with the Institute for Communications Engineering, Technische Universität München, 80333 Munich, Germany (e-mail: gerhard.kramer@tum.de).

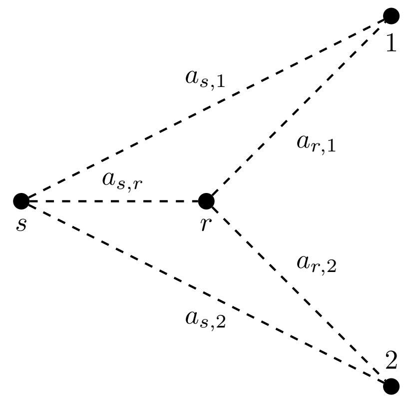


Fig. 2: AWGN-MRC with two destinations.

and where the maximization is over all $X_sX_r$.

- Direct-Transmission (DT) Rate: $C \geq R_{DT}$ where

$$R_{DT} = \max\{\min_{1\leq j\leq N} I(X_s;Y_j|X_r = x^*)\} \quad (2)$$

and where the maximization is over all $x^*$ and $X_s$.

- DF Rate: $C \geq R_{DF}$ where

$$R_{DF} = \max\left\{\min_{1\leq j\leq N} \min(I(X_sX_r;Y_j), I(X_s;Y_r|X_r))\right\} \quad (3)$$

and where the maximization is over all $X_sX_r$.

- QF Rate: $C \geq R_{QF}$ where

$$R_{QF} = \max\left\{\min_{1\leq j\leq N} \min\left(I(X_sX_r;Y_j) - I(Y_r;\widehat{Y}_r|X_sX_rY_j), I(X_s;\widehat{Y}_rY_j|X_r)\right)\right\} \quad (4)$$

and where the maximization is over all $X_sX_r\hat{Y}_r$ such that $X_s$ and $X_r$ are independent and $X_s - X_rY_r - \widehat{Y}_r$ forms a Markov chain.

### C. AWGN MRC

The (real) AWGN-MRC has real channel symbols and

$$Y_r = a_{s,r}X_s + Z_r \quad (5)$$

$$Y_j = a_{s,j}X_s + a_{r,j}X_r + Z_j \quad (6)$$

where $j \in \mathcal{T}$. The $a_{s,r}$, $a_{s,j}$, and $a_{r,j}$ are channel gains between the nodes (see Fig. 2). We later relate these gains to distances between the nodes. The $Z_r$ and $Z_j$, $j = 1, 2, \ldots, N$, are independent and identically distributed Gaussian random variables with zero mean and unit variance. We may alternatively write (5) and (6) in vector form as

$$\mathbf{Y}_j = \mathbf{A}_j\mathbf{X} + \mathbf{Z}_j \quad (7)$$

where $\mathbf{X} = (X_s\ X_r)^T$, $\mathbf{Y}_j = (Y_r\ Y_j)^T$, $\mathbf{Z} = (Z_r\ Z_j)^T$, and

$$\mathbf{A}_j = \begin{pmatrix} a_{s,r} & 0 \\ a_{s,j} & a_{r,j} \end{pmatrix}. \quad (8)$$

We consider individual average block power constraints

$$\mathrm{E}\left[\sum_{k=1}^{n} X_{s,k}^2\right] \leq nP_s, \quad \mathrm{E}\left[\sum_{k=1}^{n} X_{r,k}^2\right] \leq nP_r. \quad (9)$$

The SNR and the capacity of the link from node $u$ (with transmit power $P_u$) to node $v$ are the respective

$$\mathsf{SNR}_{u,v} = a_{u,v}^2 P_u \quad (10)$$

$$\mathsf{C}(\mathsf{SNR}_{u,v}) = \frac{1}{2}\log\left(1 + \mathsf{SNR}_{u,v}\right). \quad (11)$$

We simplify the above rate bounds for the AWGN-MRC.

- Cut-Set Bound:

$$R_{CS} = \max_{\rho}\left[\min_{1\leq j\leq N} \min\left(\mathsf{C}(\mathsf{SNR}_{s,j} + \mathsf{SNR}_{r,j} + 2\rho\sqrt{\mathsf{SNR}_{s,j}\mathsf{SNR}_{r,j}}\right), \mathsf{C}((1-\rho^2)(\mathsf{SNR}_{s,j} + \mathsf{SNR}_{s,r}))\Big)\right] \quad (12)$$

where $\rho = \mathrm{E}[X_sX_r] / \sqrt{\mathrm{E}[X_s^2]\,\mathrm{E}[X_r^2]}$ satisfies $|\rho| \leq 1$. One can restrict attention to non-negative $\rho$.

- DT Rate:

$$R_{DT} = \min_{1\leq j\leq N} \mathsf{C}(\mathsf{SNR}_{s,j}). \quad (13)$$

- DF Rate:

$$R_{DF} = \max_{\rho}\left[\min_{1\leq j\leq N} \min\left(\mathsf{C}(\mathsf{SNR}_{s,j} + \mathsf{SNR}_{r,j} + 2\rho\sqrt{\mathsf{SNR}_{s,j}\mathsf{SNR}_{r,j}}), \mathsf{C}((1-\rho^2)\mathsf{SNR}_{s,r})\right)\right]. \quad (14)$$

One can again restrict attention to non-negative $\rho$.

- QF Rate: Optimizing $X_sX_r\widehat{Y}_r$ seems difficult. Instead, we choose $X_s$ and $X_r$ to be zero-mean Gaussian with variances $P_s$ and $P_r$, respectively. We further choose $\widehat{Y}_r = Y_r + Z_r$ where $Z_r$ is zero-mean Gaussian with variance $N_r$. Optimizing $\mathrm{N}_r$ gives (see [7, p. 336-337])

$$\widetilde{R}_{QF} = \min_{1\leq j\leq N} \mathsf{C}\left(\mathsf{SNR}_{s,j} + \frac{\mathsf{SNR}_{r,j}\mathsf{SNR}_{s,r}}{\mathsf{SNR}_{s,j} + \mathsf{SNR}_{r,j} + \mathsf{SNR}_{s,r} + 1}\right). \quad (15)$$

## III. Quasi-Concavity in SNRs and $\rho$

### A. Cut-Set Bound

We consider two characterizations of $R_{CS}$. First, let $\mathbf{a}_j^T = (a_{s,j}\ a_{r,j})$ be the second row of $\mathbf{A}_j$, let $\mathbf{Q}_\mathbf{X}$ be the covariance matrix of $\mathbf{X}$ (see Appendix B), and let $\det \mathbf{M}$ be the determinant of the square matrix $\mathbf{M}$. The cut-set rate (12) can be expressed as the maximum of

$$R_{CS}(\mathbf{Q}_\mathbf{X}) = \min_{1\leq j\leq N} \min\left(\frac{1}{2}\log\left(\mathbf{a}_j^T\,\mathbf{Q}_\mathbf{X}\,\mathbf{a}_j + 1\right), \frac{1}{2}\log\left(\frac{\det \mathbf{Q}_{(\mathbf{Y}_j^T\ X_r)^T}}{P_r}\right)\right) \quad (16)$$

over the convex set of $\mathbf{Q}_\mathbf{X}$ with diagonal entries $P_s$ and $P_r$. The first logarithm in (16) is clearly concave in $\mathbf{Q}_\mathbf{X}$. The second logarithm is concave in $\mathbf{Q}_{(\mathbf{Y}_j^T\ X_r)^T}$ (see Appendix B)

and $\mathbf{Q}_{(\mathbf{Y}_j^T\ X_r)^T}$ is linear in $\mathbf{Q}_\mathbf{X}$. To prove the latter claim, observe that

$$\mathbf{Q}_{(\mathbf{Y}_j^T\ X_r)^T} = \tilde{\mathbf{A}}_j \mathbf{Q}_\mathbf{X} \tilde{\mathbf{A}}_j^T + \begin{pmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} \tag{17}$$

where $\tilde{\mathbf{A}}_j^T = \left(\mathbf{A}_j^T\ [0\ 1]^T\right)$ and $\mathbf{I}_2$ is the $2 \times 2$ identity matrix. Hence $R_{CS}(\mathbf{Q}_\mathbf{X})$ is concave in (the convex set of) $\mathbf{Q}_\mathbf{X}$ because it is the minimum of $2N$ concave functions.

Suppose next that we wish to consider $\rho$ and the SNRs individually rather than via $\mathbf{Q}_\mathbf{X}$. Define the vector

$$\mathbf{S} = (\mathsf{SNR}_{s,r}, \mathsf{SNR}_{s,1}, \cdots, \mathsf{SNR}_{s,N}, \mathsf{SNR}_{r,1}, \cdots, \mathsf{SNR}_{r,N})$$

and the functions

$$f_j(\rho, \mathbf{S}) = \mathsf{SNR}_{s,j} + \mathsf{SNR}_{r,j} + 2\rho\sqrt{\mathsf{SNR}_{s,j}\mathsf{SNR}_{r,j}} \tag{18}$$

$$g_j(\rho, \mathbf{S}) = (1-\rho^2)\left(\mathsf{SNR}_{s,j} + \mathsf{SNR}_{s,r}\right) \tag{19}$$

$$R_{CS}(\rho, \mathbf{S}) = \min_{1 \le j \le N} \min\left(\mathsf{C}(f_j(\rho, \mathbf{S})), \mathsf{C}(g_j(\rho, \mathbf{S}))\right). \tag{20}$$

We establish the following results.

*Lemma 1:* $f_j(\rho, \mathbf{S})$ and $g_j(\rho, \mathbf{S})$ are concave in $\rho$, concave in $\mathbf{S}$, and quasi-concave in $(\rho^2, \mathbf{S})$ for $0 \le \rho^2 \le 1$ and non-negative $\mathbf{S}$.

*Proof:* Concavity with respect to $\rho$ is established by observing that $f_j(\rho, \mathbf{S})$ is linear in $\rho$, and $g_j(\rho, \mathbf{S})$ is linear in $-\rho^2$ which is concave in $\rho$.

Consider next concavity with respect to $\mathbf{S}$. The Hessian of $f_j(\rho, \mathbf{S})$ with respect to $\mathbf{S}$ has only one non-zero eigenvalue

$$-\frac{\rho}{2} \cdot \frac{\mathsf{SNR}_{s,j}^2 + \mathsf{SNR}_{r,j}^2}{\mathsf{SNR}_{s,j}^{3/2}\mathsf{SNR}_{r,j}^{3/2}}. \tag{21}$$

Thus, $f_j(\rho, \mathbf{S})$ is concave in $\mathbf{S}$. The function $g_j(\rho, \mathbf{S})$ is linear in $\mathbf{S}$, and thus concave in $\mathbf{S}$.

Now consider quasi-concavity with respect to $(\rho^2, \mathbf{S})$. Substituting $a = \mathsf{SNR}_{s,j}, b = \mathsf{SNR}_{r,j}, c = \rho^2$ into the fifth function of Lemma 9, we find that $f_j(\rho, \mathbf{S})$ is quasi-concave in $(\rho^2, \mathbf{S})$. For the $g_j(\rho, \mathbf{S})$, observe that $ab$ is quasi-concave for non-negative $(a, b)$ (see the first function of Lemma 9). This implies

$$(\lambda a_1 + \bar{\lambda} a_2)(\lambda b_1 + \bar{\lambda} b_2) \ge \min(a_1 b_1, a_2 b_2) \tag{22}$$

for $0 \le \lambda \le 1$, and where $\bar{\lambda} = 1 - \lambda$. Substituting $a_i = 1 - \rho_i^2$ and $b_i = \mathsf{SNR}_{s,j,i} + \mathsf{SNR}_{s,r,i}$ for $i = 1, 2$, we find that $g_j(\rho, \mathbf{S})$ is quasi-concave in $(\rho^2, \mathbf{S})$. ■

*Theorem 1:* $R_{CS}(\rho, \mathbf{S})$ is concave in $\rho$, concave in $\mathbf{S}$, and quasi-concave in $(\rho^2, \mathbf{S})$ for $0 \le \rho^2 \le 1$ and non-negative $\mathbf{S}$.

*Proof:* $R_{CS}(\rho, \mathbf{S})$ involves taking logarithms and minima of (quasi-) concave functions. The results thus follow by applying Lemma 1 and Lemma 8, Parts 2 and 3. ■

*Corollary 1:* Consider $\mathbf{S}$ as a function of the powers $\mathbf{P} = (P_s, P_r)$. Then $R_{CS}(\rho, \mathbf{S}(\mathbf{P}))$ is quasi-concave in $(\rho^2, \mathbf{P})$ for $0 \le \rho^2 \le 1$ and non-negative $\mathbf{P}$.

*Proof:* The proof follows from Theorem 1 and because $\mathbf{S}$ is a linear function of $\mathbf{P}$. ■

### B. DF Rate

Consider the functions

$$g_j^*(\rho, \mathbf{S}) = (1-\rho^2)\mathsf{SNR}_{s,r} \tag{23}$$

$$R_{DF}(\rho, \mathbf{S}) = \min_{1 \le j \le N} \min\left(\mathsf{C}(f_j(\rho, \mathbf{S})), \mathsf{C}(g_j^*(\rho, \mathbf{S}))\right). \tag{24}$$

*Theorem 2:* $R_{DF}(\rho, \mathbf{S})$ is concave in $\rho$, concave in $\mathbf{S}$, and quasi-concave in $(\rho^2, \mathbf{S})$ for $0 \le \rho^2 \le 1$ and non-negative $\mathbf{S}$.

*Proof:* The proof is similar to that of Theorem 1. ■

*Corollary 2:* $R_{DF}(\rho, \mathbf{S}(\mathbf{P}))$ is quasi-concave in $(\rho^2, \mathbf{P})$ for $0 \le \rho^2 \le 1$ and non-negative $\mathbf{P}$.

*Proof:* See the proof of Corollary 1. ■

### C. DT Rate

The DT rate (13) is clearly concave in $\mathbf{S}$ and $\mathbf{P}$.

### D. QF Rate

Consider the functions

$$h_j(\mathbf{S}) = \mathsf{SNR}_{s,j} + \frac{\mathsf{SNR}_{r,j}\mathsf{SNR}_{s,r}}{\mathsf{SNR}_{s,j} + \mathsf{SNR}_{r,j} + \mathsf{SNR}_{s,r} + 1} \tag{25}$$

$$\widetilde{R}_{QF}(\mathbf{S}) = \min_{1 \le j \le N} \mathsf{C}(h_j(\mathbf{S})). \tag{26}$$

We establish the following results.

*Lemma 2:* $h_j(\mathbf{S})$ is quasi-concave in $(\mathsf{SNR}_{r,j}, \mathsf{SNR}_{s,r})$ for non-negative SNRs.

*Proof:* Substitute $a = \mathsf{SNR}_{r,j}, b = \mathsf{SNR}_{s,r}, k = \mathsf{SNR}_{s,j} + 1$ into the second function of Lemma 9, and apply Lemma 8, Part 1. ■

*Theorem 3:* $\widetilde{R}_{QF}(\mathbf{S})$ is quasi-concave in the non-negative SNRs if the $\mathsf{SNR}_{s,j}$, $j = 1, 2, \ldots, n$, are held fixed.

*Proof:* Apply Lemmas 2 and Lemma 8, Parts 2 and 3. ■

## IV. Quasi-Concavity in Relay Position

Suppose the channel gain for the node pair $(i, j)$ is

$$a_{i,j} = \sqrt{\xi_{i,j}} \Big/ D_{i,j}^{\alpha/2} \tag{27}$$

where $\xi_{i,j}$ is a "fading" gain, $D_{i,j} = \|\mathbf{i} - \mathbf{j}\|$ is the Euclidean distance between the positions $\mathbf{i}$ and $\mathbf{j}$ of nodes $i$ and $j$, respectively, and $\alpha \ge 2$ is a path-loss exponent. We thus have

$$\mathsf{SNR}_{i,j} = \frac{\xi_{i,j} P_i}{D_{i,j}^\alpha} = \frac{\xi_{i,j} P_i}{\|\mathbf{i} - \mathbf{j}\|^\alpha}.$$

We establish quasi-concavity results in $\rho^2$ and $\mathbf{r}$.

### A. Cut-Set Bound

Consider the functions (18)-(20) but relabeled as $f_j(\rho, \mathbf{r})$, $g_j(\rho, \mathbf{r})$, and $R_{CS}(\rho, \mathbf{r})$ to emphasize the dependence on the considered parameters.

*Lemma 3:* $f_j(\rho, \mathbf{r})$ and $g_j(\rho, \mathbf{r})$ are quasi-concave in $\mathbf{r}$ for fixed $\rho$, $0 \le \rho \le 1$. Furthermore, $f_j(\rho, \mathbf{r})$ is quasi-concave in $(\rho^2, \mathbf{r})$ for $0 \le \rho^2 \le 1$.

*Proof:* Consider the functions

$$\tilde{f}_j(\rho, D^\alpha) = \frac{\xi_{s,j} P_s}{D_{s,j}^\alpha} + \frac{\xi_{r,j} P_r}{D^\alpha} + 2\rho\sqrt{\frac{\xi_{s,j} P_s}{D_{s,j}^\alpha}\frac{\xi_{r,j} P_r}{D^\alpha}} \tag{28}$$

$$\tilde{g}_j(\rho, D^\alpha) = (1-\rho^2)\left(\frac{\xi_{s,j}P_s}{D^\alpha_{s,j}} + \frac{\xi_{s,r}P_s}{D^\alpha}\right) \tag{29}$$

which are quasi-linear in $D^\alpha$ for fixed $\rho$ since they are decreasing in $D^\alpha$. But $D^\alpha_{r,j}$ and $D^\alpha_{s,r}$ are convex functions of $\mathbf{r}$ for $\alpha \geq 1$, and thus Lemma 8, Part 5, establishes that $f_j(\rho, \mathbf{r})$ and $g_j(\rho, \mathbf{r})$ are quasi-concave in $\mathbf{r}$ for fixed $\rho$.

Next, substitute $a = D^\alpha$ and $b = \rho^2$ into the third function of Lemma 9, and use Lemma 8, Part 1, to show that $\tilde{f}_j(\rho, D^\alpha)$ is quasi-concave in $(\rho^2, D^\alpha)$. But $\tilde{f}_j$ is decreasing in $D^\alpha$ and $D^\alpha_{r,j}$ is convex in $\mathbf{r}$, so Lemma 8, Part 5, establishes that $f_j(\rho, \mathbf{r})$ is quasi-concave in $(\rho^2, \mathbf{r})$. ■

Unfortunately, $\tilde{g}_j$ is quasi-*convex* (and not quasi-concave) in $(\rho^2, D^\alpha)$. To see this, substitute $a = D^\alpha$ and $b = \rho^2$ into the fourth function of Lemma 9. Quasi-concavity would have been useful since it would have permitted using Lemma 8, Parts 2 and 4, to establish the quasi-concavity of

$$R_{CS}(\mathbf{r}) = \max_\rho \left[\min_{1\leq j\leq N} \min\left(\mathsf{C}(f_j(\rho, \mathbf{r})), \mathsf{C}(g_j(\rho, \mathbf{r}))\right)\right]. \tag{30}$$

Based on our numerical results, $R_{CS}(\mathbf{r})$ seems to be quasi-concave in $\mathbf{r}$ but we have been unable to prove this. Nevertheless, Lemma 3 suffices to establish an intermediate result which is useful in Sec. V when we study $\rho = 0$.

*Theorem 4:* $R_{CS}(\rho, \mathbf{r})$ is quasi-concave in $\mathbf{r}$ for fixed $\rho$, $0 \leq \rho \leq 1$.

*Proof:* $R_{CS}(\rho, \mathbf{r})$ is the minimum of functions that are quasi-concave in $\mathbf{r}$. Lemma 8, Part 2, thus establishes the theorem. ■

### B. DF Rate

The quasi-convexity of $\tilde{g}_j(\rho, D^\alpha)$ relaxes for the DF rate (24). Consider the negative of the fourth function of Lemma 9 with $k_1 = 0$:

$$f(a, b) = (1-b)k_2/a. \tag{31}$$

This function is quasi-linear in $(a, b)$ since both its superlevel and sublevel sets are convex. This result implies the following theorem. We again consider the functions (23)-(24) but relabeled as $g^*_j(\rho, \mathbf{r})$ and $R_{DF}(\rho, \mathbf{r})$. We further define

$$\tilde{g}^*_j(\rho, D^\alpha) = (1-\rho^2)\frac{\xi_{s,r}P_s}{D^\alpha} \tag{32}$$

$$R_{DF}(\mathbf{r}) = \max_\rho \left[\min_{1\leq j\leq N} \min\left(\mathsf{C}(f_j(\rho, \mathbf{r})), \mathsf{C}(g^*_j(\rho, \mathbf{r}))\right)\right]. \tag{33}$$

*Theorem 5:* $R_{DF}(\rho, \mathbf{r})$ is quasi-concave in $(\rho^2, \mathbf{r})$ for $0 \leq \rho^2 \leq 1$, and $R_{DF}(\mathbf{r})$ is quasi-concave in $\mathbf{r}$.

*Proof:* $\tilde{g}^*_j(\rho, D^\alpha)$ is quasi-linear in $(\rho^2, D^\alpha)$ and decreasing in $D^\alpha$. Furthermore, $D^\alpha_{s,r}$ is convex in $\mathbf{r}$, and thus Lemma 8, Part 5, establishes that $g^*_j(\rho, \mathbf{r})$ is quasi-concave in $(\rho^2, \mathbf{r})$. $R_{DF}(\rho, \mathbf{r})$ is therefore quasi-concave in $\mathbf{r}$, as it is the minimum of quasi-concave functions (see Lemma 8, Part 2). Furthermore, $R_{DF}(\mathbf{r})$ is concave in $\mathbf{r}$ by Lemma 8, Part 4. ■

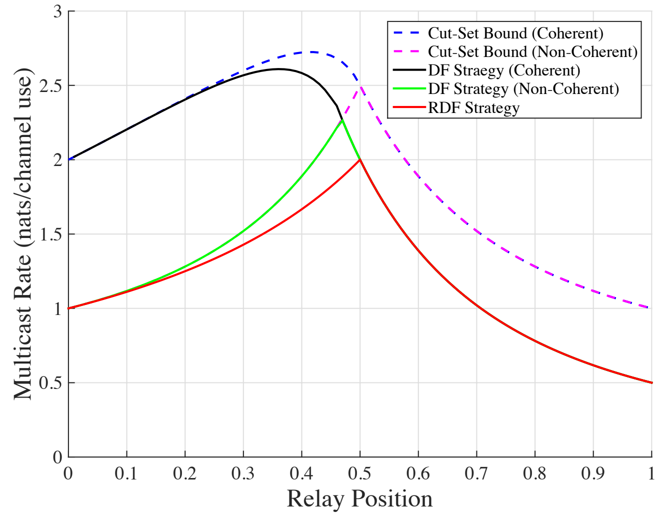


Fig. 3: Relay channel rates for low SNR and $P = 1$.

## V. DF Performance

This section presents numerical results for the DF strategy and compares them to results from our previous work [3]–[5]. We consider 1-, 2-, and 3-dimensional MRCs with different numbers $N$ of destination nodes. For simplicity, we consider the low-SNR or broadband regime where

$$\mathsf{C}(\mathsf{SNR}) = \frac{1}{2}\log(1 + \mathsf{SNR}) \to \frac{1}{2}\mathsf{SNR}. \tag{34}$$

In other words, we consider the cut-set and DF rates without the logarithms. This approach is valid not only in the limit of low SNR, but more generally because we proved our quasi-concavity results *without* taking logarithms. Furthermore, in the low-SNR regime the rates of full-duplex and half-duplex transmission are the same under a block power constraint.

We choose $P_s = P_r = P = 1$, $\alpha = 2$, and $\xi_{u,v} = 1$ for all node pairs $(u, v)$. We study both coherent transmission where $\rho$ is optimized and non-coherent transmission with $\rho = 0$. The rates are in nats/channel use. Alternatively, suppose we use sinc pulses sampled at $2W$ samples per second, where $W$ is the (one-sided) signal bandwidth. Suppose further that the (one-sided) noise power spectral density is 1 Watt/Hz. Then at low SNR the rates in nats/channel use are the same as the rates in nats/sec.

### A. One Dimension

Consider a relay channel ($N = 1$) where the source is at the origin ($\mathbf{s} = 0$) and the destination is at point 1 ($\mathbf{1} = 1$). Fig. 3 shows the low-SNR cut-set bounds, DF rates, and the routing-based DF (RDF) rates developed in [3]. For example, the non-coherent rates with $\rho = 0$ are

$$R_{DF} \to \frac{P}{2}\min\left(\frac{\xi_{s,1}}{\|\mathbf{s}-\mathbf{1}\|^\alpha} + \frac{\xi_{r,1}}{\|\mathbf{r}-\mathbf{1}\|^\alpha}, \frac{\xi_{s,r}}{\|\mathbf{s}-\mathbf{r}\|^\alpha}\right) \tag{35}$$

$$R_{RDF} \to \max_{0\leq\beta\leq1} \frac{P}{2}\left[\min\left(\frac{\xi_{r,1}}{\|\mathbf{r}-\mathbf{1}\|^\alpha}, \frac{\xi_{s,r}\beta}{\|\mathbf{s}-\mathbf{r}\|^\alpha}\right) + \frac{\xi_{s,1}(1-\beta)}{\|\mathbf{s}-\mathbf{1}\|^\alpha}\right]. \tag{36}$$

Observe that all curves are quasi-concave (but not concave) in $\mathbf{r}$. Theorems 4 and 5 predict the quasi-concavity for all curves except for the coherent cut-set bound.

The best relay positions for the two strategies are different. For example, $\mathbf{r} = 0.5$ maximizes $R_{RDF}$ while the $\mathbf{r}$ maximizing $R_{DF}$ is closer to the source. This is because when the source transmits, the relay *and* the destination listen and the destination "collects" information. The relay can thus be positioned closer to the source while maintaining the same information rate from the source to the relay, and from the source-relay pair to the destination. At the optimal positions, we compute $R_{DF} \approx 2.25P$ nats/sec and $R_{RDF} = 2P$ nats/sec, so the DF gain is $\approx 12\%$. We remark that the non-coherent cut-set bound coincides with (the non-coherent) $R_{DF}$ and $R_{RDF}$ for a large range of $\mathbf{r}$.

### B. Two Dimensions

Consider $N = 5$ destinations positioned on a square in the two-dimensional Euclidean plane with the source node at the origin. Fig. 4a plots the non-coherent $R_{DF}$ as a function of the relay position. The best relay position $\mathbf{r}^*_{DF}$ is shown by a circle and the corresponding rate is $R_{DF} \approx 0.011P$ nats/sec. Fig. 4c plots the low-SNR two-hop rate

$$R_{2H} \to \min_{1\le j\le 5} \frac{P}{2} \min\left(\frac{\xi_{s,r}}{\|\mathbf{s}-\mathbf{r}\|^\alpha}, \frac{\xi_{r,j}}{\|\mathbf{r}-\mathbf{j}\|^\alpha}\right) \tag{37}$$

as a function of the relay position. The best relay position $\mathbf{r}^*_{2H}$ is shown by a circle and the corresponding two-hop rate is $R_{2H} \approx 0.0095P$ nats/sec. The non-coherent DF gain is thus $\approx 12\%$.

Fig. 4b and Fig. 4d show contour plots for $R_{DF}$ and $R_{2H}$. The contours form convex regions, as predicted by Theorem 5. Again, the relay position maximizing $R_{DF}$ lies closer to the source than the relay position maximizing $R_{2H}$.

### C. Three Dimensions

Consider $N = 5$ destinations positioned in 3-dimensional Euclidean space as in Fig. 5. The figure also shows the convex hull (a polyhedron) of the points. The points $\mathbf{r}^*_{\mathrm{DF}}$ and $\mathbf{r}^*_{\mathrm{2H}}$ denote the relay positions that maximize the non-coherent $R_{DF}$ and $R_{2H}$, respectively. We remark that $\mathbf{r}^*_{\mathrm{DF}}$ and $\mathbf{r}^*_{\mathrm{2H}}$ remain unchanged if more destinations are positioned inside the polyhedron. This is because the points in the polyhedron can receive at least the same rate as the worst of the five nodes at the corner points.

## VI. Discussion

We considered real AWGN channels. For complex AWGN channels, quasi-concavity in $\mathbf{r}$ will likely not be valid because the *phases* of the channel gains $a_{u,v}$ will change with $\mathbf{r}$, and this means that $\rho$ must be adjusted differently for different destination nodes. We remark that this effect is "local" in the sense that for large carrier frequencies the phase variations are sensitive to changes in $\mathbf{r}$. A pragmatic approach would then be to optimize $\mathbf{r}$ for non-coherent transmission ($\rho = 0$) even if beam-forming is permitted.

## VII. Conclusion

Various quasi-concavity results were established for AWGN MRCs. Quasi-concavity lets one use efficient optimization algorithms to maximize the rate and to place the relay.

## Appendix A
## Concave and Quasi-Concave Functions

### A. Definitions

Consider the following sets. The *domain* $\mathcal{D}_f$ of a real-valued function $f : \mathbf{R}^n \to \mathbf{R}$ is the set of arguments for which $f$ is defined. The *hypograph* and *hypergraph* of $f$ are the respective

$$\mathcal{H}_f = \{(\mathbf{x}, y)\,|y \le f(\mathbf{x})\} \tag{38}$$

$$\widehat{\mathcal{H}}_f = \{(\mathbf{x}, y)\,|y \ge f(\mathbf{x})\}. \tag{39}$$

The *superlevel* and *sublevel* sets of $f$ with respect to $\beta \in \mathbf{R}$ are the respective

$$\mathcal{S}_{f,\beta} = \{\mathbf{x}|f(\mathbf{x}) \ge \beta\} \tag{40}$$

$$\widehat{\mathcal{S}}_{f,\beta} = \{\mathbf{x}|f(\mathbf{x}) \le \beta\}. \tag{41}$$

Concave and quasi-concave functions can be defined via the convexity of these sets. Recall that a set $\mathcal{S}$, $\mathcal{S} \subseteq \mathbf{R}^n$, is *convex* if for any two points $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\mathcal{S}$ and for any $\lambda$ satisfying $0 \le \lambda \le 1$ we have

$$\lambda\mathbf{x}_1 + \bar{\lambda}\mathbf{x}_2 \in \mathcal{S} \tag{42}$$

where $\bar{\lambda} = 1 - \lambda$. Suppose that $\mathcal{D}_f$ is convex. The function $f$ is *concave* over $\mathcal{D}_f$ if and only if its hypograph $\mathcal{H}_f$ is convex. Similarly, $f$ is *convex* over $\mathcal{D}_f$ if and only if $\widehat{\mathcal{H}}_f$ is convex. The function $f$ is *quasi-concave* over $\mathcal{D}_f$ if and only if all its superlevel sets are convex, and $f$ is *quasi-convex* over $\mathcal{D}_f$ if and only if all its sublevel sets are convex. A function that is quasi-convex and quasi-concave is called *quasi-linear*. For example, any non-increasing or non-decreasing function is quasi-linear.

### B. Basic Properties

Two properties of concave and quasi-concave functions are as follows; these properties are often used as the definitions of such functions. Similar properties exist for convex and quasi-convex functions.

*Lemma 4:* The function $f$ is concave if and only if

$$f(\lambda\mathbf{x}_1 + \bar{\lambda}\mathbf{x}_2) \ge \lambda f(\mathbf{x}_1) + \bar{\lambda} f(\mathbf{x}_2) \tag{43}$$

for all $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\mathcal{D}_f$ and for all $0 \le \lambda \le 1$.

*Lemma 5:* The function $f$ is quasi-concave if and only if

$$f(\lambda\mathbf{x}_1 + \bar{\lambda}\mathbf{x}_2) \ge \min(f(\mathbf{x}_1), f(\mathbf{x}_2)) \tag{44}$$

for all $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\mathcal{D}_f$ and for all $0 \le \lambda \le 1$.

The next two properties assume that $f$ is twice differentiable and that $\mathcal{D}_f$ is convex. Let $\mathbf{H}_f(\mathbf{x})$ and $\mathbf{B}_f(\mathbf{x})$ be the respective Hessian and bordered Hessian of $f$ at $\mathbf{x}$.

*Lemma 6:* (see [8, Sec. 3.1.4]) $f$ is concave if and only if $\mathbf{H}_f(\mathbf{x})$ is negative semidefinite for all $\mathbf{x} \in \mathcal{D}_f$.

*Lemma 7:* (see [9, p. 771]) $f$ is quasi-concave on the open and convex set $\mathcal{D}_f$ if the determinants $D_2, D_3, \ldots, D_n$ of the respective second to $n$th leading principal minors of $\mathbf{B}_f(\mathbf{x})$ satisfy $(-1)^k D_k < 0$ for $k = 2, 3, \ldots, n$ and for all $\mathbf{x} \in \mathcal{D}_f$.

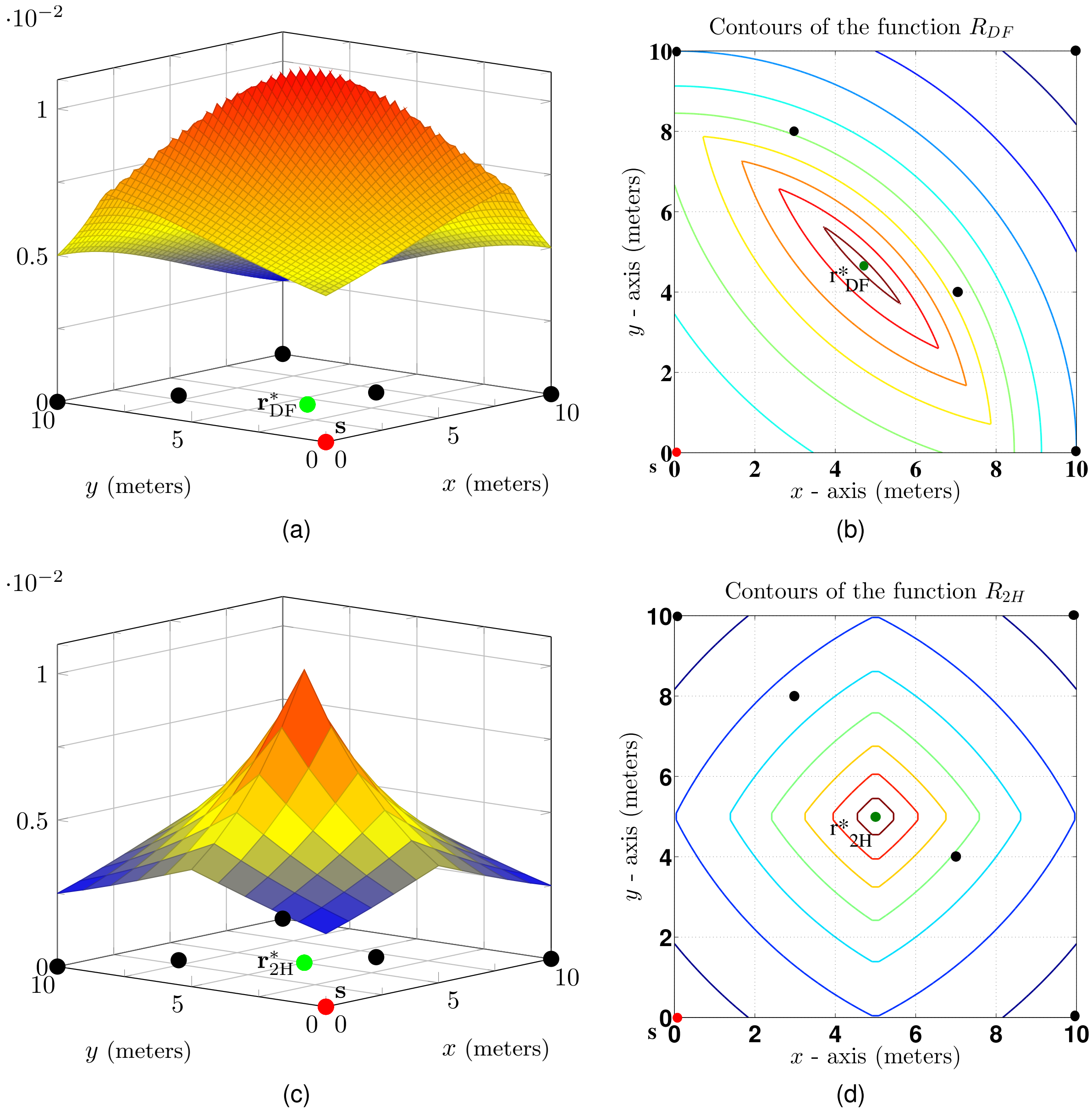


Fig. 4: (a) $R_{DF}$ for $N = 5$; (b) Contour plot of $R_{DF}$; (c) $R_{2H}$ for the same network; (d) Contour plot of $R_{2H}$.

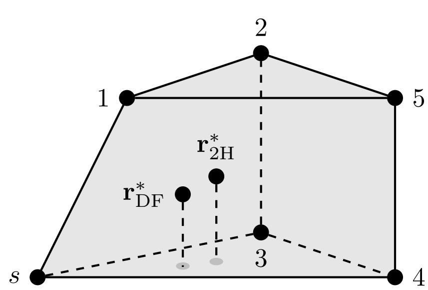


Fig. 5: $N = 5$ destination geometry in three dimensions.

### C. Compositions Preserving Quasi-Concavity

The following compositions preserve quasi-concavity.

*Lemma 8:* Suppose $f$ and $f_i$, $1 \le i \le n$, are quasi-concave, then so are the functions

1) $h = k_1 f + k_2$, where $k_1 \ge 0$ and $k_2 \in \mathbf{R}$;
2) $h = \min_{1\le i\le n} f_i$;
3) $h = g \circ f$ where $f$ is quasi-concave and $g$ is non-decreasing;
4) $h(\mathbf{a}) = \sup_{\mathbf{b}\in\mathcal{B}} f(\mathbf{a},\mathbf{b})$ where $\mathcal{B}$ is a convex set;
5) $h(\mathbf{a},\mathbf{b}) = f(g(\mathbf{a}),\mathbf{b})$ where $g$ is convex and $f(\tilde{\mathbf{a}},\mathbf{b})$ is non-increasing in $\tilde{\mathbf{a}}$ for fixed $\mathbf{b}$.

*Proof:* Properties 1)-4) are standard (see [8, Sec. 3.4]). For property 5), observe that

$$\begin{aligned} & h(\lambda\mathbf{a}_1 + \bar{\lambda}\mathbf{a}_2, \lambda\mathbf{b}_1 + \bar{\lambda}\mathbf{b}_2) \\ & = f(g(\lambda\mathbf{a}_1 + \bar{\lambda}\mathbf{a}_2), \lambda\mathbf{b}_1 + \bar{\lambda}\mathbf{b}_2) \\ & \overset{(a)}{\ge} f(\lambda g(\mathbf{a}_1) + \bar{\lambda} g(\mathbf{a}_2), \lambda\mathbf{b}_1 + \bar{\lambda}\mathbf{b}_2) \\ & \overset{(b)}{\ge} \min\left(f(g(\mathbf{a}_1),\mathbf{b}_1), f(g(\mathbf{a}_2),\mathbf{b}_2)\right) \end{aligned} \tag{45}$$

where $(a)$ follows because $g(\lambda\mathbf{a}_1 + \bar{\lambda}\mathbf{a}_2) \le \lambda g(\mathbf{a}_1) + \bar{\lambda} g(\mathbf{a}_2)$ and $f(\tilde{\mathbf{a}},\mathbf{b})$ is non-increasing in $\tilde{\mathbf{a}}$. Step $(b)$ follows because $f$ is quasi-concave. ∎

### D. Examples of Quasi-Concave Functions

We establish quasi-concavity for several useful functions.

*Lemma 9:* The following functions are quasi-concave for $\mathbf{x} = (a\ b\ c)$ with non-negative entries.

1) $f(\mathbf{x}) = ab$
2) $f(\mathbf{x}) = \frac{ab}{a+b+k}$ for a positive constant $k$
3) $f(\mathbf{x}) = k_1/a + 2\sqrt{k_2 b/a}$ for positive constants $k_1, k_2$
4) $f(\mathbf{x}) = -(1-b)(k_1 + k_2/a)$ for positive constants $k_1, k_2$
5) $f(\mathbf{x}) = a + b + 2\sqrt{abc}$

*Proof:* We use bordered Hessians $\mathbf{B}_f(\mathbf{x})$ and the derivatives $D_k$ of their $k$th leading principal minors, $k = 2, 3, \ldots, n$. The results extend to non-negative $\mathbf{x}$ by using continuity at zero values, except for the third and fourth functions where $a = 0$ makes the functions undefined.

1) We have $D_2 < 0$ and $D_3 > 0$ for
$$\mathbf{B}_f(\mathbf{x}) = \begin{pmatrix} 0 & bc & ac \\ bc & 0 & c \\ ac & c & 0 \end{pmatrix}.$$
2) We have $D_2 < 0$ and $D_3 > 0$ for
$$\mathbf{B}_f(\mathbf{x}) = \begin{pmatrix} 0 & \frac{b(b+k)}{(a+b+k)^2} & \frac{a(a+k)}{(a+b+k)^2} \\ \frac{b(b+k)}{(a+b+k)^2} & \frac{-2b(b+k)}{(a+b+k)^3} & \frac{2ab+(a+b+k)k}{(a+b+k)^3} \\ \frac{a(a+k)}{(a+b+k)^2} & \frac{2ab+(a+b+k)k}{(a+b+k)^3} & \frac{-2a(a+k)}{(a+b+k)^3} \end{pmatrix}.$$
3) We have $D_2 < 0$ and $D_3 > 0$ for
$$\mathbf{B}_f(\mathbf{x}) = \begin{pmatrix} 0 & -\frac{k_1+\sqrt{k_2 ab}}{a^2} & \sqrt{\frac{k_2}{ab}} \\ -\frac{k_1+\sqrt{k_2 ab}}{a^2} & \frac{4k_1+3\sqrt{k_2 ab}}{2a^3} & -\frac{\sqrt{k_2}}{2a^{3/2}\sqrt{b}} \\ \sqrt{\frac{k_2}{ab}} & -\frac{\sqrt{k_2}}{2a^{3/2}\sqrt{b}} & -\frac{\sqrt{k_2}}{2b^{3/2}\sqrt{a}} \end{pmatrix}.$$
4) We have $D_2 < 0$ and $D_3 > 0$ for
$$\mathbf{B}_f(\mathbf{x}) = \begin{pmatrix} 0 & \frac{(1-b)k_2}{a^2} & k_1 + \frac{k_2}{a} \\ \frac{(1-b)k_2}{a^2} & -\frac{2(1-b)k_2}{a^3} & -\frac{k_2}{a^2} \\ k_1 + \frac{k_2}{a} & -\frac{k_2}{a^2} & 0 \end{pmatrix}.$$
5) We have $D_2 < 0$, $D_3 > 0$ and $D_4 < 0$ for
$$\mathbf{B}_f(\mathbf{x}) = \begin{pmatrix} 0 & 1+\sqrt{\frac{bc}{a}} & 1+\sqrt{\frac{ac}{b}} & \sqrt{\frac{ab}{c}} \\ 1+\sqrt{\frac{bc}{a}} & -\frac{\sqrt{bc}}{2a^{3/2}} & \frac{1}{2}\sqrt{\frac{c}{ab}} & \frac{1}{2}\sqrt{\frac{b}{ac}} \\ 1+\sqrt{\frac{ac}{b}} & \frac{1}{2}\sqrt{\frac{c}{ab}} & -\frac{\sqrt{ac}}{2b^{3/2}} & \frac{1}{2}\sqrt{\frac{a}{bc}} \\ \sqrt{\frac{ab}{c}} & \frac{1}{2}\sqrt{\frac{b}{ac}} & \frac{1}{2}\sqrt{\frac{a}{bc}} & -\frac{\sqrt{ab}}{2c^{3/2}} \end{pmatrix}. \tag{46}$$

∎

## Appendix B
## Covariance Matrices and Concavity

The covariance matrix of a real-valued random column vector $\mathbf{V}$ is

$$\mathbf{Q}_{\mathbf{V}} = \mathrm{E}\left[(\mathbf{V} - \mathrm{E}[\mathbf{V}])(\mathbf{V} - \mathrm{E}[\mathbf{V}])^T\right]. \tag{47}$$

A useful property of covariance matrices is as follows (see [10, p. 684]). If $\mathbf{Q}_{\mathbf{V}}^*$ is a principal minor of $\mathbf{Q}_{\mathbf{V}}$, then the function

$$f(\mathbf{Q}_{\mathbf{V}}) = \log \frac{\det \mathbf{Q}_{\mathbf{V}}}{\det \mathbf{Q}_{\mathbf{V}}^*} \tag{48}$$

is concave in $\mathbf{Q}_{\mathbf{V}}$.

## Acknowledgment

M. Thakur and G. Kramer were supported by the German Ministry of Education and Research in the framework of an Alexander von Humboldt Professorship.

## References

[1] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inf. Theory*, vol. 51, no. 9, Sep. 2005.
[2] V. Aggarwal, A. Bennatan, and A. R. Calderbank, “On maximizing coverage in Gaussian relay channels,” *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2518–2536, Jun. 2009.
[3] M. Thakur, N. Fawaz, and M. Médard, “Optimal relay location and power allocation for low-SNR broadcast relay channels,” *IEEE Int. Conf. Comp. Commun., INFOCOM 2011, Shanghai, China*, Apr. 2011.
[4] ——, “On the geometry of wireless network multicast in 2-D,” *IEEE Int. Symp. Inf. Theory, St. Petersburg, Russia*, Jul. 2011.
[5] ——, “Reducibility of joint relay positioning and flow optimization problem,” *IEEE Int. Symp. Inf. Theory, Boston, MA*, Jul. 2012.
[6] T. M. Cover and A. El Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572-584, Sep. 1979.
[7] G. Kramer, I. Maric, and R. D. Yates, “Cooperative Communications,” *Foundations and Trends in Networking*, vol. 1, nos. 3-4, pp. 271-425, 2006.
[8] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.
[9] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming*. Hoboken: Wiley, 2006.
[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 2nd edition, 2006.